DOCUMENT RESUME

ED 362 526 TM 020 512

AUTHOR Schattgen, Sharon Ford; And Others
TITLE Using Criterion-Referenced Tests To Evaluate Chapter 1 Programs: A Lesson from Missouri.
PUB DATE Apr 92
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Tests; *Compensatory Education; *Criterion Referenced Tests; Elementary Secondary Education; Evaluation Criteria; *Federal Programs; Grade 1; Information Networks; Mastery Tests; Mathematics Tests; *Program Evaluation; Reading Tests; State Programs; Test Construction; Testing Programs; *Test Use
IDENTIFIERS *Education Consolidation Improvement Act Chapter 1; Michigan First Grade Reading and Math Test; *Missouri; Missouri Mastery and Achievement Tests

ABSTRACT

The Missouri Department of Elementary and Secondary Education uses state-developed, criterion-referenced tests for the purpose of evaluating programs funded by Chapter 1 of the Elementary and Secondary Education Act. The agency's efforts to obtain approval from the U.S. Department of Education (USDE) to use state norms for evaluation purposes are described, and evidence is presented to demonstrate that the state tests meet the USDE criteria for approval. The second section of the paper describes the innovative and efficient processes that Missouri uses to collect, aggregate, and report evaluation data to the local education agencies. The Missouri Mastery and Achievement Tests and the First-Grade Reading and Mathematics Test, the two tests developed for Chapter 1 evaluation, are appropriate and efficient for program evaluation throughout most of the grade-level continuum. Both meet or exceed the USDE requirements. Electronic transmission of information, a feature of the Missouri system, from collection of pretest scores to the generation and communication of data and results saves time and effort for local Chapter 1 personnel. Four appendixes provide sample reports generated by the testing process and a local school report form. Eight tables present evaluation findings. (Contains 7 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
SHARON F. SCHATTGEN
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 362 526

Using Criterion-Referenced Tests to Evaluate Chapter 1 Programs: A Lesson from Missouri

Sharon Ford Schattgen
Center for Educational Assessment, University of Missouri-Columbia

Shelly Lock
Missouri Department of Elementary and Secondary Education

Don Snyder
Missouri Department of Elementary and Secondary Education

Paper Presented as Part of a Symposium Entitled
"Chapter 1 Evaluation Leadership in State Departments of Education"
at the Annual Meeting of the
American Educational Research Association

San Francisco, California, April 1992

TM020512

Abstract

This paper explains how a state education agency (SEA) uses state-developed, criterion-referenced achievement tests to evaluate Chapter 1 programs. The paper describes the SEA's efforts to obtain approval from the United States Department of Education (USDE) to use state norms for evaluation purposes, and it presents evidence demonstrating that the state tests meet the USDE criteria for approval. The paper also describes the innovative reports and efficient procedures that the SEA employs to collect, aggregate, and report evaluation data to local school districts.

Using Criterion-Referenced Tests to Evaluate Chapter 1 Programs: A Lesson from Missouri

The purpose of this paper is to explain how a state education agency, the Missouri Department of Elementary and Secondary Education (MDESE), uses state-developed, criterion-referenced tests for the purpose of evaluating programs funded by Chapter 1 of the Elementary and Secondary Education Act. The paper is divided into two sections. The first section presents evidence demonstrating that the Missouri tests meet the criteria established by the United States Department of Education (USDE) for the use of state norms in the evaluation of Chapter 1 programs. The second section describes the innovative and efficient processes that MDESE uses to collect, aggregate, and report evaluation data to local education agencies (LEAs).

Use of State-Normed, Criterion-Referenced Tests In Evaluation

Description of State Tests

The *Missouri Mastery and Achievement Tests (MMAT)* represent the cornerstone of Missouri's statewide assessment program. The *MMAT* battery measures educational objectives, referred to as Key Skills, and consists of 34 criterion-referenced tests assessing student performance in grades 2 through 10. Tests for grade 2 cover Reading, Language Arts, and Mathematics. Tests for grades 3 through 10 cover Reading, Language Arts, Mathematics, Science, and Social Studies. At each grade, there are at least two equivalent test forms. The *First-Grade Reading and Mathematics Test (First-Grade Test)*, also a criterion-referenced instrument, measures educational objectives in Reading, Language Arts, and Mathematics that are linked to the grade 2 Key Skills measured by the *MMAT*. Eventually, there will be two equivalent forms of the *First-Grade Test*, although at the present time only one form is available.

State regulations, based on the Missouri Excellence in Education Act of 1985 (House Bill 463), require all school districts to administer the *MMAT* battery or another approved measure of the Key Skills each year to students in two non-consecutive levels in the grade span 2 through 6 and to students in two non-consecutive levels in the grade span 7 through 10. The battery is administered in the spring, so that it can be used for summative evaluation purposes, although districts may also administer it in the fall.

The *MMAT* battery and the *First-Grade Test* may be appropriately used to evaluate the effectiveness of Chapter 1 programs throughout most of the grade-level continuum. Corresponding subject tests are linked vertically across all grades on the basis of the common core of content measured. The MDESE utilizes a spring-to-spring evaluation cycle for Chapter 1 programs. At grades 2 through 10, the *MMAT* may be administered to end-of-year

students in order to obtain pretest scores to compare with posttest data obtained in the following spring. At grade 1, the *First-Grade Test* may be administered to end-of-year students in order to obtain pretest scores to compare with grade 2 *MMAT* posttest scores obtained in the following spring. (School districts are not required to administer the *First-Grade Test*, but they are encouraged to utilize it in the identification of students who are eligible for Chapter 1 services. If districts are using the *MMAT* at grade 2 to evaluate the effectiveness of Chapter 1 programs, they must use the *First-Grade Test* to obtain pretest data.)

From 1987 through 1990, the *MMAT* yielded Chapter 1-usable scores referred to as estimated comparable national percentile ranks. These estimated norms were obtained by equating raw scores on *MMAT* subtests to raw scores on corresponding subtests from a nationally normed achievement battery. This equating procedure, described in detail by Schattgen and Osterlind (1988), allowed districts to engage in dual-purpose testing and to allocate resources to the administration of the *MMAT*, which, as a criterion-referenced battery, has more instructional utility than its norm-referenced counterparts. The results of studies investigating the technical properties of the estimated comparable national percentile ranks suggested that they were valid for making eligibility decisions (Schattgen & Osterlind, 1989) but were not as reliable as might be desired (Schattgen, 1990b). Consequently, the MDESE was pleased when federal regulations were modified in order to allow the use of local (state or district) norms in Chapter 1 evaluations (USDE, 1990).

USDE Criteria for Approval of State Norms

The USDE (1990) stipulates that local rather than national norms may be used for Chapter 1 evaluation purposes, if the achievement test used meets certain criteria. The agency identifies (p. 9) "two areas of concern in developing state norms for use in evaluation: the validity of the state tests for measuring typical Chapter 1 instructional objectives," referred to as "content validity," and the "representativeness of the norms." The salient points of the criteria associated with the respective areas are given below.

Content Validity

According to USDE (1990) requirements governing the use of state norms for Chapter 1 evaluation purposes, the state must show that

> each level of its test is a valid measure of the subject matter relevant to Chapter 1 evaluation, namely reading, mathematics and other language arts content typically measured by total tests and subtests of norm-referenced achievement tests used in the evaluation of Chapter 1 programs. Content validity must be demonstrated using the correlation between the state test and any nationally normed achievement test. . . . Evidence of adequate correlation would be a

> correlation between total scores of the state test and the selected nationally normed test of at least 0.75, and at least 0.70 for subtests. (p. 9)

The USDE regulations postulate that correlations should be computed on a sample of at least 200.

Representativeness of Norms

In order to be in compliance with USDE policies (1990) regarding the use of state norms for Chapter 1 evaluation purposes, the state must present evidence that the

> norms developed are representative of the achievement levels of all students in the population to be tested. Such evidence would include a description of the probability sampling design employed and a comparison of the relevant characteristics of the students who constitute the norms with the characteristics of the student population. . . . In order to ensure that state norms are quite stable, the sample size for any one grade level should be no less than 2, 000. . . . (p. 9)
>
> Pretest-posttest gains calculated for Chapter 1 students that are based on state norms should be consistent with the NCE gains for these students based on national norms. For each level of the state test, a subsample of at least 200 students should be administered a nationally normed test. The standard deviation of the NCE scores on the nationally normed test should be calculated. If it is different from 21.06, the pretest-posttest gains should be expressed in standard deviation units (using the state norming distribution) and then multiplied by 21.06. (p. 10)

MDESE Application for Approval of State Norms

The MDESE filed an application (Schattgen, 1990a) requesting the USDE's approval to use the state norms yielded by the *MMAT* and the *First-Grade Test* for the purpose of evaluating Chapter 1 programs. The application and the addendum to it presented evidence, which is given below, demonstrating how the *MMAT* and the *First-Grade Test* meet the USDE (1990) criteria for approval of state norms.

Content Validity of State Tests

The *MMAT* and the *First-Grade Test* demonstrate acceptable content validity as it was defined in 1990 by the USDE. This conclusion is based on the analysis of data resulting from concurrent administrations of the two

Missouri tests and two nationally normed batteries to representative samples of Missouri students in Spring 1990. The nationally normed tests administered were the *Iowa Tests of Basic Skills (ITBS)* at grades 1 through 8 and the *Tests of Achievement and Proficiency (TAP)* at grades 9 and 10.

Pearson product-moment correlation coefficients indicating the relationships between corresponding subject-test raw-score distributions yielded by the *MMAT* and the *ITBS* at grades 2 through 8 and the *MMAT* and the *TAP* at grades 9 and 10 are shown in Table 1. Pearson product-moment correlation coefficients indicating the relationships between corresponding subject-test raw-score distributions yielded by the *First-Grade Test* and the grade-1 *ITBS* are also shown in Table 1. The coefficients were computed using data from subsamples of the probability sample on which the test was normed (see description of norming sample below). The number of examinees taking each pair of tests (the *n* used to compute each correlation) is given in the table. Note that the *n* for each subsample exceeds 200, the *n* required for approval.

The coefficients indicate that, in general, there are strong relationships between subject-test raw scores on the *MMAT/First-Grade Test* and the *ITBS/TAP*. At all grades and subjects except grade-1 reading and language arts, the coefficients exceed the criterion of .75 needed for approval. At .734, the grade-1 reading coefficient is very close to the criterion; at .636, the grade-1 language arts coefficient is slightly lower than the criterion.

Representativeness of State Norms

Types reported.

For the purpose of conducting Chapter 1 program evaluations, both the *MMAT* and the *First-Grade Test* yield a state percentile rank and a corresponding normal curve equivalent (NCE) score in each Chapter 1 area (Reading Comprehension, Language Arts, and Mathematics). Within Mathematics, a state percentile and a corresponding NCE are reported for basic skills (computation) and also for advanced skills (application and problem solving). The percentiles and NCEs are within-group norms reflecting an examinee's performance relative to the sample on which each test is annually normed.

Representativeness.

The norming samples for the *MMAT* and the *First-Grade Test* are groups of examinees representative of the Missouri student population. At each grade, students in the sample are selected for inclusion in the sample using a probability design, specifically the stratified-random-cluster technique. Two stratification variables are used to determine the characteristics of the sample: school-district size and school-district type. School buildings within the districts represent the clusters.

Table 1

Pearson Product-Moment Coefficients for Corresponding First-Grade Test/MMAT and ITBS/TAP Subjects, Spring 1990 Administration

	Grade									
Subject	1	2	3	4	5	6	7	8	9[a]	10[b]
Reading	.734	.870	.773	.828	.814	.784	.806	.776	.754	.805
	562	1772	1188	1252	788	1169	1050	975	593	1294
Language Arts	.636	.767	.761	.797	.792	.830	.826	.814	—	—
	248	1704	1124	1121	1307	1154	1223	1152	—	—
Mathematics	.749	.748	.809	.807	.848	.850	.863	.886	.854	.832
	589	1793	1215	1269	1145	1192	887	1320	1383	1238

Note. The *n* is presented below each coefficient.

[a]Language Arts is not tested at grade 9.

[b]Language Arts is not tested at grade 10.

There are 11 sampling categories based on the stratification variables. Table 2 presents a description of each category. In each grade in each category, a building has a random chance of being selected for inclusion in the sample. Table 3 shows the proportion of the Missouri student population in each grade in each category.

In order to annually norm the *MMAT* and the *First-Grade Test*, a new sample is drawn for each spring administration. Approximately one-tenth of the total student population is selected for inclusion in the total sample. This proportion is larger than necessary in order to compensate for attrition in the sample due to unavoidable circumstances (e.g., classes are not scheduled to be in session at the building during the specific week that the sample is to be tested).

Table 4 presents the number of students in each grade in each category selected for inclusion in the Spring 1990 sample. (The number in each cell represents the corresponding proportion of the population that is shown in Table 3.) As previously noted, not all students selected for inclusion in the sample are actually tested as part of the sample due to a variety of unavoidable circumstances. However, the actual number of students tested in the sample in each grade, also shown in Table 4, exceeds 2000, the number required for approval.

Tables 5 through 7 present *MMAT* and *First-Grade Test* data for the Spring 1990 correlation and norming samples and for the Spring 1990 population. Raw-score means and standard deviations of the reading tests, language arts tests, and mathematics tests are shown in Tables 5, 6, and 7, respectively. These data provide evidence of the representativeness of the correlation and norming samples because sample and population values for corresponding grades and subjects are very similar. Note that the number of examinees in each correlation and norming subsample exceeds the number required for approval (200 and 2000, respectively).

Table 8 presents data from the Spring 1990 correlation sample for the nationally normed tests used to validate the *MMAT* and the *First-Grade Reading Test*. Standard deviations of NCE score distributions for the reading, language arts, and mathematics subtests of the *Iowa Tests of Basic Skills (ITBS)* and for the reading and mathematics subtests of the *Tests of Achievement and Proficiency (TAP)* are shown. Most of the the *ITBS/TAP* NCE standard deviations are comparable to the national standard of 21.06, although only one value (grade-8 mathematics) equals 21.0. (The standard deviations range from 17.2 to 21.6; the median standard deviation is 18.95, and the mean standard deviation is 19.2.) Consequently, the MDESE will adjust NCE gains based on the state norms for all grade-level subject tests except grade-8 mathematics in accordance with how the corresponding *ITBS/TAP* NCE standard deviation differs from 21.06.

Table 2

Sampling Categories

Category	Description
A	western urban (Kansas City)
B	eastern urban (St. Louis)
C	southwestern urban (Springfield)
D	out-state, mid-sized urban; university-dominated economy
E	out-state, mid-sized urban; industry-dominated economy
F	extreme southeastern rural ("boot-heel" area); agriculture-based economy
G	northern rural; agriculture-based economy
H	southwestern rural; agriculture-based economy
I	southeastern rural; timber/mining-based economy
J	western suburban (Kansas City suburbs)
K	eastern suburban (St. Louis suburbs)

Table 3

Proportion of Student Population in Each Cell

Grade

Category	Grade 1	2	3	4	5	6	7	8	9	10
A	.049	.050	.049	.048	.047	.045	.044	.042	.039	.037
B	.070	.062	.059	.056	.051	.049	.046	.045	.072	.048
C	.027	.030	.030	.028	.030	.030	.028	.030	.027	.029
D	.039	.038	.037	.038	.038	.037	.038	.037	.036	.041
E	.045	.043	.043	.044	.044	.043	.046	.043	.041	.042
F	.040	.040	.041	.041	.044	.045	.044	.044	.044	.042
G	.125	.126	.127	.129	.128	.131	.128	.127	.123	.124
H	.072	.073	.072	.075	.073	.076	.072	.076	.074	.077
I	.221	.226	.228	.227	.231	.230	.234	.234	.224	.228
J	.200	.198	.201	.200	.200	.200	.206	.206	.215	.221
K	.112	.114	.113	.114	.114	.114	.114	.116	.105	.111

Table 4

Number of Students Selected for Inclusion and Number of Students Tested in Sample, Spring 1990 Administration

	Grade									
Category	1	2	3	4	5	6	7	8	9	10
A	321	324	319	310	305	292	288	268	250	243
B	454	400	382	362	328	318	301	293	466	311
C	175	193	194	184	192	192	184	193	173	191
D	252	250	240	247	244	242	246	242	235	266
E	289	280	282	288	288	281	298	280	266	273
F	261	259	266	266	285	295	286	286	286	270
G	814	816	824	836	834	850	830	826	800	804
H	465	477	470	488	476	493	470	494	479	499
I	1438	1465	1477	1474	1504	1492	1516	1523	1457	1481
J	1300	1289	1308	1297	1298	1296	1337	1340	1398	1434
K	725	741	732	743	741	744	738	750	685	722
Total Selected	6494	6494	6494	6495	6495	6495	6494	6495	6495	6494
Total Tested	4411	4604	5018	4865	4762	4673	4209	5160	3921	4693

Table 5

Raw-Score Statistics for First-Grade Test/MMAT Reading Tests, Spring 1990 Administration

	Correlation Sample		Norming Sample		Population	
Grade	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1	30.8	6.6	32.0	5.9	31.5	6.2
	562		4,411		28,394	
2	24.2	6.6	25.1	6.0	25.5	5.7
	1,772		4,604		42,674	
3	40.5	9.4	42.6	8.3	42.7	8.2
	1,188		5,018		51,121	
4	34.6	10.8	38.7	9.4	39.0	9.2
	1,252		4,865		44,136	
5	34.9	8.4	35.8	7.9	35.9	7.8
	788		4,762		44,245	
6	38.3	9.1	41.1	7.7	40.7	8.1
	1,169		4,673		49,221	
7	39.4	11.2	40.8	10.9	41.9	10.6
	1,050		4,209		40,007	
8	42.5	10.3	43.8	10.0	43.9	9.8
	975		5,160		47,941	
9	41.9	10.3	43.3	9.8	43.5	10.0
	593		3,921		34,853	
10	44.1	9.0	44.2	8.8	43.8	9.0
	1,294		4,693		44,254	

Note. The *n*/*N* is presented below each set of statistics.

Table 6

Raw-Score Statistics for First-Grade Test/MMAT Language Arts Tests, Spring 1990 Administration

	Correlation Sample[a]		Norming Sample		Population	
Grade	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1	21.2	3.8	20.8	4.0	20.4	4.2
	248		4,411		28,394	
2	30.5	9.2	32.3	8.5	33.0	8.2
	1,704		4,604		42,674	
3	30.7	7.4	31.8	6.4	31.7	6.5
	1,124		5,018		51,121	
4	24.6	7.5	25.9	6.9	26.2	6.9
	1,121		4,865		44,136	
5	28.6	6.6	30.2	5.7	30.4	5.6
	1,307		4,762		44,245	
6	30.5	8.3	32.2	7.4	31.6	7.6
	1,154		4,673		49,221	
7	35.7	9.1	36.9	8.6	37.7	8.4
	1,223		4,209		40,007	
8	37.7	10.5	38.6	9.9	38.4	9.9
	1,152		5,160		47,941	
9			37.8	9.5	37.9	9.6
			3,921		34,853	
10			37.1	9.4	36.7	9.4
			4,693		44,254	

Note. The *n*/*N* is presented below each set of statistics.

[a] *TAP* Language Arts subtest was not administered to correlation sample at grades 9 and 10.

Table 7

Raw-Score Statistics for First-Grade Test/MMAT Mathematics Tests, Spring 1990 Administration

	Correlation Sample		Norming Sample		Population	
Grade	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1	54.6	4.9	54.7	5.0	54.1	5.5
	589		4,411		28,249	
2	25.9	5.5	26.6	4.9	26.8	4.7
	1,793		4,604		42,628	
3	56.5	9.9	57.8	8.8	57.6	8.9
	1,215		5,018		51,181	
4	39.5	8.5	40.6	7.6	41.3	7.4
	1,269		4,865		44,265	
5	38.0	11.5	40.8	10.8	41.0	10.7
	1,145		4,762		44,202	
6	74.3	19.1	77.6	16.7	75.1	17.7
	1,192		4,673		49,254	
7	50.4	18.1	54.4	18.2	56.9	18.0
	887		4,209		39,951	
8	62.0	20.7	66.2	18.4	65.9	18.2
	1,320		5,160		47,998	
9	38.0	13.2	39.8	12.7	40.3	13.1
	1,383		3,921		34,909	
10	57.3	17.7	57.7	17.0	57.2	16.9
	1,238		4,693		44,332	

Note. The *n*/*N* is presented below each set of statistics.

Table 8

Standard Deviations of Normal Curve Equivalent Scores for ITBS/TAP Tests, Spring 1990 Administration

	Subject		
Grade	Reading	Language Arts[a]	Mathematics
1	19.8	19.7	18.7
	562	248	589
2	20.7	19.0	20.2
	1,772	1,704	1,793
3	18.3	18.5	20.2
	1,188	1,124	1,215
4	17.2	18.8	18.5
	1,252	1,121	1,269
5	17.7	19.4	20.0
	788	1,307	1,145
6	17.4	18.6	20.1
	1,169	1,154	1,192
7	18.0	18.4	18.5
	1,050	1,223	887
8	18.9	20.1	21.0
	975	1,152	1,320
9	19.4	—	18.6
	593		1,383
10	21.6	—	20.0
	1,294		1,238

Note. The correlation sample *n* is presented below each standard deviation.

[a] *TAP* Language Arts subtest was not administered to correlation sample at grades 9 and 10.

Collection and Aggregation of Evaluation Data and Reporting of Results

Using MMAT for Chapter 1 Purposes

The MDESE encourages all Chapter 1 programs in Missouri to coordinate their testing efforts with regular classroom and district programs. To facilitate this coordination, the Chapter 1 office at the MDESE realigned evaluation guidelines and procedures to allow local education agencies to use the *MMAT* battery, which is criterion-referenced and measures the Key Skills, for Chapter 1 evaluation purposes. The spring-to-spring annual evaluation cycle allows for a single test administration for selection and/or evaluation purposes, reduces the testing burden on schools districts, and provides accurate and timely information on the statewide Chapter 1 population and its relationship to the total student population. (See MDESE, 1991a, for a complete description of policies and guidelines that apply to Chapter 1 programs in Missouri.)

MMAT Chapter 1 Reports

To promote the use of *MMAT* results for Chapter 1 purposes, the MDESE offers two reports: The Chapter 1 Eligibility List and the Chapter 1 Pre/Post Evaluation Report. These reports, depicted in Appendices A and B, respectively, are described below.

Eligibility list.

The Chapter 1 Eligibility List shows all students in a grade in a building who are eligible, based on *MMAT* results, to receive Chapter 1 services. The standards for eligibility that are applied to students taking the *MMAT* in the spring reflect the grade placement for the next school year during which services are to be received. For each Chapter 1 program area, a student's eligibility status as well as his/her state percentile rank (referred to as the Chapter 1 eligibility rank) and corresponding NCE are shown. Students are listed in ascending order according to the Reading percentile rank.

Note that this report shows performance in basic and advanced skills within Mathematics. (In Mathematics, basic skills involve computation, while advanced skills involve application and problem solving.) In Reading, advanced skills involve comprehension and are represented by the score reported. A score is not reported for basic skills in Reading, and scores are not reported for basic and advanced skills in Language Arts.

A dotted line appears after the name of the last student eligible for services in Reading. Students listed below the dotted line may be eligible for services in Language Arts and/or Mathematics but not Reading. There is a space on the right side of the report to indicate a change in a student's status for receiving Chapter 1 services. The unduplicated number of educationally deprived students listed on the report appears at the end of the list.

Evaluation report.

The Chapter 1 Pre/Post Evaluation Report presents evaluation data for individual students that are ultimately aggregated to determine program effectiveness. This report lists all students designated on *MMAT* answer sheets as formally identified Chapter 1 participants having a previous year's *MMAT* score in the program area(s) in which they received services during the current academic year. For each student, the report presents the previous year's *MMAT* pretest score (which was entered on the answer sheet at the time of testing and scanned during the scoring process), the current year's *MMAT* posttest score, and the difference or gain score achieved by the student in each subject in which s/he received services. Pretest and posttest scores are presented as state percentile ranks and NCEs; gain scores are presented as NCEs.

Aggregating and Reporting Evaluation Results to LEAs

If there are inaccuracies in the *MMAT* data presented on the Pre/Post Evaluation Report, Chapter 1 personnel at the LEA return a corrected copy of the report to the Center for Educational Assessment at the University of Missouri-Columbia no later than July 1. On behalf of MDESE, the Center aggregates the evaluation data and provides feedback to the LEA via the MDESE Chapter 1 Program Impact Report. (Districts using a standardized test other than the *MMAT* for Chapter 1 evaluation, or using the *MMAT* but not participating in the optional Chapter 1 Pre/Post Evaluation Report, submit results on scannable forms.)

Program Impact Report

The MDESE Chapter 1 Program Impact Report indicates whether a building is identified for program improvement based on aggregate performance (MDESE, 1991b). Two types of Program Impact reports, building and district, are generated, and they are formatted identically. This report, shown in Appendix C, is divided into four major sections.

- The first section contains descriptive information regarding student characteristics. The Chapter 1 student membership is reported by public and non-public schools. (A participant is a student who received any amount of Chapter 1 instruction.) The number of students who were reported as having either a pretest score or a posttest score, or both is shown in the column labeled 'Total Reported/Tested.' The numbers of students tested who were designated as having an IEP (Individualized Education Program) or being of LEP (Limited English Proficiency) are also shown. Aggregated totals for each category are presented in the bottom row.

• The Evaluation Results section shows the number of students with both pretest and posttest scores. The mean NCE and percentile for the pretest, the mean NCE and percentile for the posttest, and the mean and median NCE gain are shown for each of the following subject areas and skill levels: Overall and Advanced Reading; Overall, Advanced, and Basic Mathematics; and Overall Language Arts. Aggregated data for each category are presented in the bottom row.

• The Program Improvement section repeats the total number of students for whom there are both pretest and posttest scores and indicates whether the building or district is targeted for Program Improvement by the mean or the median NCE gain. The district may choose which statistic (mean or median NCE gain) to use to determine the effectiveness of the Chapter 1 program, but that choice must be applied to all of the district's Chapter 1 buildings.

• The last section lists: Grade level; Test Battery, Pretest and Posttest; Subtest; Test Level; Norming Dates; Test Dates; Days Between the Norming Dates and Test Dates; Net Days Difference Pretest and Posttest. If more than one test, subtest, or level is used within a grade in a school building, this fact is noted by separate lines within each grade. The last column of this section shows the percent of student retention from pretest to posttest. This is one index of student mobility within Chapter 1 classes. A large number in this column indicates low mobility since students pre-tested early in the program remained through the posttest date.

Targeting LEAs for Program Improvement

Chapter 1 programs are evaluated based on standardized test scores and progress toward meeting the desired outcomes specified in the Chapter 1 application. In Missouri, a building is targeted for program improvement if there is not an aggregate gain of at least +1.0 NCE in basic and/or advanced skills (MDESE, 1991b). A building could also be targeted for program improvement if it fails to show progress toward meeting desired outcomes. Therefore, it is important for school districts to develop appropriate desired outcomes that can be easily measured and to select an appropriate test for use in evaluation.

A local annual review must be conducted by the LEA to determine if standardized test aggregate performance requirements were met and if substantial progress toward meeting desired outcomes was made. This review determines whether a program improvement plan needs to be developed. Following the review, the district completes the Attachment A form, depicted in Appendix D, and submits it to the MDESE. The MDESE then determines whether the district must begin program improvement

activities. (See MDESE, 1991b, for a complete explanation of Missouri's program improvement plan.)

Summary

The *MMAT* and the *First-Grade Test* are criterion-referenced achievement batteries which may be appropriately and efficiently used to evaluate the effectiveness of Chapter 1 programs throughout most of the grade-level continuum. Both instruments meet and/or exceed the criteria established by the USDE (1990) for approval of state norms for use in Chapter 1 program evaluations. Furthermore, innovative reports and procedures facilitate the convenient use of *MMAT* results for Chapter 1 purposes. The electronic transmission of information—from the collection of the pretest scores on the student answer sheets to the generation and communication of evaluation data and results—saves Chapter 1 personnel at the LEAs time and effort.

References

Missouri Department of Elementary and Secondary Education. (1991a). *Chapter 1, ESEA, manual of operational policies and guidelines.* Jefferson City, MO: Author.

Missouri Department of Elementary and Secondary Education. (1991b). *Chapter 1 Missouri state plan for program improvement.* Jefferson City, MO: Author.

Schattgen, S. F. (1990a). *An application for approval of state norms for the purpose of evaluating chapter 1 programs.* Columbia, MO: Missouri Department of Elementary and Secondary Education and University of Missouri.

Schattgen, S. F. (1990b). The reliability of norm-referenced information obtained from an objective-referenced test. (Doctoral dissertation, University of Missouri-Columbia, 1990). *Dissertation Abstracts International, 51,* 3714A.

Schattgen, S. F., & Osterlind, S. J. (1988, April). *Estimating norm-referenced information from a criterion-referenced test: An application of the ORT only model.* Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA. (ERIC Document Reproduction Service No. 295 972)

Schattgen, S. F., & Osterlind, S. J. (1989, March). *The validity of norm-referenced information obtained from an objective-referenced test using the ORT only model.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

United States Department of Education. (1990). *Guidelines and criteria for equating and norming studies conducted for chapter 1 evaluation.* Washington, DC: Author.

APPENDIX A

MISSOURI MASTERY AND ACHIEVEMENT TESTS
CHAPTER 1 ELIGIBILITY LIST

Building: PLUM RIPE ELEMENTARY
District: GOOD SOIL RURAL DISTRICT #1
Code: 000-000-0000

Subject:
Grade: 4
Form: 4 / E
Test Date: SPRING 1991

This roster ranks students in ascending order according to the Chapter 1 Eligibility Rank achieved on the reading cluster. The corresponding NCE (Normal Curve Equivalent) score is also displayed. In addition to the reading cluster score, language arts and mathematics scores are reported, as are scores for advanced and basic skills within math. (Eligibility in math is determined by the overall score.) The reading cluster score represents performance in advanced reading skills. A dotted line appears after the name of the last student eligible by the reading criterion.

These ranks represent state norms and are intended for use in determining eligibility for Chapter 1 services received after July 1, 1991, and for obtaining pretest scores for Chapter 1 reporting for the 1991–92 school year.

Legend for ethnic heritage identification: A = American Indian or Alaskan Native B = Asian or Pacific Islander C = Hispanic D = Black E = White

	Demographic Data				Eligible in			READING		LANGUAGE		MATHEMATICS					
												Overall		Basic Skills		Advanced Skills	
NAME	IEP/LEP	Birth Date	Gender	Ethnic Heritage	R	L	M	Chapter 1 Eligibility Rank	(NCE)	Chapter 1 Eligibility Rank	(NCE)	Chapter 1 Eligibility Rank	(NCE)	Chapter 1 Eligibility Rank	(NCE)	Chapter 1 Eligibility Rank	(NCE)	+ or –
	IEP		M	E	R	L	M	3	(10)	5	(15)	27	(37)	16	(29)	32	(40)	
	LEP		M	E	R	L	M	7	(19)	15	(28)	21	(33)	15	(28)	46	(49)	
	BOTH		F	E	R	L		14	(27)	1*	(1)	43	(46)	22	(34)	56	(54)	
	IEP		M	E	R	L	M	22	(34)	8	(20)	24	(35)	35	(42)	15	(28)	
	LEP		M	E	R	L	M	26	(37)	16	(29)	24	(35)	16	(29)	27	(37)	
			M	E	R	L	M	26	(37)	30	(39)	24	(35)	26	(37)	24	(35)	
			M	D	R		M	28	(38)	53	(52)	40	(45)	48	(49)	27	(37)	
			F	E	R	L	M	32	(40)	15	(28)	24	(35)	26	(37)	22	(34)	
			F	E	R	L	M	32	(40)	35	(42)	24	(35)	14	(27)	40	(27)	
			M	E	R			32	(40)	43	(46)	40	(45)	43	(46)	43	(46)	
			M	E	R	L		40	(45)	35	(42)	51	(51)	61	(56)	35	(42)	
			M	E	R			40	(45)	48	(49)	43	(46)	58	(54)	40	(45)	
			F	E				43	(46)	48	(49)	43	(46)	53	(52)	35	(42)	
			F	E				45	(47)	45	(47)	51	(51)	43	(46)	58	(54)	
			M	E		L		48	(49)	27	(37)	43	(46)	51	(51)	40	(45)	
			F	E				48	(49)	53	(52)	58	(54)	43	(46)	48	(49)	

* Student completed less than 70% of items

MISSOURI MASTERY AND ACHIEVEMENT TESTS

CHAPTER 1 PRE/POST EVALUATION REPORT

Building: PLUM RIPE ELEMENTARY SCHOOL
District: GOOD SOIL RURAL DISTRICT #1
Code: 000-000-0000

Subject:
Grade: 3
Form: 3/C
Test Date: SPRING 1991

APPENDIX B

This report presents Chapter 1 evaluation data which will be transmitted to the Missouri Department of Elementary and Secondary Education. These data are based on the pretest and posttest scores of students identified as ESEA Chapter 1 participants on Spring 1991 answer sheets. A student's gain in each subject is computed using the NCE achieved on the MMAT in Spring 1990 (derived from the state percentile rank that was reported on the Spring 1991 answer sheet) and the NCE achieved on the MMAT in Spring 1991. Pretest and posttest scores are presented as CHAPTER 1 ELIGIBILITY RANKS and corresponding NCEs; gain scores are presented as NCEs. In addition to reporting scores, this roster indicates each student's IEP or LEP status, birth year, gender, and ethnic heritage.

Legend for ethnic heritage identification: A = American Indian or Alaskan Native B = Asian or Pacific Islander C = Hispanic D = Black E = White

NAME	Demographic Data				READING			LANGUAGE			MATH (OVERALL)			MATH (BASIC)			MATH (ADVANCED)		
	IEP LEP Both	Birth Year	Gender	Ethnic Heritage	PRETEST Chapter 1 Eligibility Rank/NCE	POSTTEST Chapter 1 Eligibility Rank/NCE	GAIN NCE	PRETEST Chapter 1 Eligibility Rank/NCE	POSTTEST Chapter 1 Eligibility Rank/NCE	GAIN NCE	PRETEST Chapter 1 Eligibility Rank/NCE	POSTTEST Chapter 1 Eligibility Rank/NCE	GAIN NCE	PRETEST Chapter 1 Eligibility Rank/NCE	POSTTEST Chapter 1 Eligibility Rank/NCE	GAIN NCE	PRETEST Chapter 1 Eligibility Rank/NCE	POSTTEST Chapter 1 Eligibility Rank/NCE	GAIN NCE
	__	82__	M__	E__	24___/35	28/38	+3												
	IEP__	83__	F__	E__							16___/29	23/34	+5	28__/28	21/33	+5	30__/30	26/37	-1
	__	82__	F__	E__	19___/32	31/40	+8	12___/25	19/32	+7									
	__	82__	F__	E__	28___/38	50/50	+12	22___/34	28/38	+4									
	BTH__	83__	F__	D__	18___/31	22/34	+3	17___/30	50/50	+20									

NOTE: This report should list data only for 1990–91 Chapter 1 students and only for 1990–91 subject programs. It should not include any students who were not in a Chapter 1 program during 1990–91. Indicate corrections on a photocopy of the report in the spaces provided and send it to CEA by July 1.

MISSOURI DEPARTMENT OF ELEMENTARY AND SECONDARY EDUCATION
CHAPTER 1, ESEA PROGRAM IMPACT
BUILDING REPORT REGULAR SCHOOL YEAR

BUILDING: PLUM RIPE ELEMENTARY SCHOOL
DISTRICT: GOOD SOIL RURAL DISTRICT #1
CODE: 000-000-0000

SUBJECT AREA: READING

APPENDIX C

DESCRIPTIVE INFORMATION						EVALUATION RESULTS							
GRADE	TOTAL PARTICIPANTS PUBLIC	TOTAL PARTICIPANTS NON-PUB	TOTAL REPORTED TESTED	IEP STUDENTS TESTED	LEP STUDENTS TESTED	TYPE SCORE	NUMBER PRE AND POST	PRETEST MEAN NCE	PRETEST MEAN %ILE	POSTTEST MEAN NCE	POSTTEST MEAN %ILE	MEAN NCE GAIN	MEDIAN NCE GAIN
2	5	0	5	1	0	Overall:	5	29.9	17	33.3	21	3.4	-1.9
						Advanced:	0	.	.	.	.	.	.
3	13	0	13	0	0	Overall:	13	27.7	14	35.0	24	7.3	5.6
						Advanced:	13	27.7	14	35.0	24	7.3	5.6
4	9	0	9	1	0	Overall:	9	34.9	24	28.1	15	-6.8	-7.9
						Advanced:	9	34.9	24	28.1	15	-6.8	-7.9
5	11	0	11	0	0	Overall:	11	28.5	15	29.8	17	1.3	1.3
						Advanced:	11	28.5	15	29.8	17	1.3	1.3
6	6	0	6	0	0	Overall:	6	23.1	10	28.8	16	5.7	5.7
						Advanced:	6	23.1	10	28.8	16	5.7	5.7
AGGREGATED TOTALS:	44	0	44	2	0	Overall:	44	29.0	16	31.2	19	2.2	2.0
						Advanced:	39	28.9	16	31.0	18	2.1	2.6

PROGRAM IMPROVEMENT IN READING FOR:	NUMBER PRE AND POST			TRIGGERED BY MEAN NCE GAIN			TRIGGERED BY MEDIAN NCE GAIN		
	Overall	Advan.	Basic	Overall	Advan.	Basic	Overall	Advan.	Basic
000-000-0000 PLUM RIPE ELEMENTARY SCHOOL	44	39		NO	NO		NO	NO	

MISSOURI DEPARTMENT OF ELEMENTARY AND SECONDARY EDUCATION
CHAPTER 1, ESEA PROGRAM IMPACT
BUILDING REPORT REGULAR SCHOOL YEAR

BUILDING: PLUM RIPE ELEMENTARY SCHOOL
DISTRICT: GOOD SOIL RURAL DISTRICT #1
CODE: 000-000-0000

SUBJECT AREA: READING

APPENDIX C (CONT'D)

GRADE	PRE- OR POST TEST	TEST BATTERY	SUBTEST	TEST LEVEL	NORM DATE	TEST DATE	DAYS BETWEEN NORM & TEST DATES	NET DAYS DIFF PRE TO POST	NUMBER TESTED Overall	NUMBER TESTED Advan.	NUMBER TESTED Basic	PERCENT RETAINED PRE TO POST (Overall)	NUMBER USED FOR EVALUATION (Overall)
2	PRE	ITBS	R-COM	7	04/30	04/02	-28		5	0			
	POST	OTHER	R-TOT	2		04/08	.	.	5	0		100%	5
3	PRE	MMAT	R-TOT	2	04/15	04/06	-9		13	13			
	POST	MMAT	R-TOT	3	04/15	04/05	-10	-1	13	13		100%	13
4	PRE	MMAT	R-TOT	3	04/15	04/06	-9		9	9			
	POST	MMAT	R-TOT	4	04/15	04/05	-10	-1	9	9		100%	9
5	PRE	MMAT	R-TOT	4	04/15	04/06	-9		11	11			
	POST	MMAT	R-TOT	5	04/15	04/05	-10	-1	11	11		100%	11
6	PRE	MMAT	R-TOT	5	04/15	04/06	-9		6	6			
	POST	MMAT	R-TOT	6	04/15	04/05	-10	-1	6	6		100%	6

DEPARTMENT OF ELEMENTARY AND SECONDARY EDUCATION
P.O. BOX 480, JEFFERSON CITY, MISSOURI 65102
CHAPTER 1, ESEA
LOCAL SCHOOL REPORT

ATTACHMENT "A"

Whether or not attainment of an average gain of one NCE is made in basic, advanced, and overall skills, Chapter 1 regulations require consideration of locally developed outcomes and local conditions before a final determination is made regarding the necessity of program improvement activities. This form provides such information to help the local school district and the state agency determine if program improvement activities must be initiated.

This form is to be completed for <u>each</u> Chapter 1 subject area. Include building level analysis for <u>each</u> building having more than 10 participating Chapter 1 students, grades PreK-12, regardless of whether a plus (+) one average aggregate NCE gain has been achieved in basic, advanced and overall skills. (If more forms are required, copies may be made.)

School District Name	District Code Number

A. Locally Developed Desired Outcomes (Must address both basic and advanced skills)

Progress toward meeting learner outcomes will be considered if local program objectives have been achieved.

Indicate locally developed desired outcomes for schools shown on page 10 of the 1990-91 project application for each eligible building and Chapter 1 subject area. (A separate copy should be submitted for PreK-K-grade 1:)

Outcome 1: _____

_____

Outcome 2: _____

_____

B. Local District Position on Program Improvement

Depending on the strength of any supporting data, the school district is asked to identify its position with respect to program improvement for each building listed on the back of this form.

POSITION 1

Based on a review of program effectiveness at participating Chapter 1 schools, the building(s) identified on the back of this form are those which, subject to Missouri Department of Elementary and Secondary Education confirmation, are in need of a Program Improvement Plan. Documentation which supports this process is maintained. The Program Improvement Plan will be kept on file and be submitted to the local board of education and to the SEA. Plans for these schools are to be submitted no later than January 15, 1992, or 60 days after official notification by the SEA.

POSITION 2

Local conditions identified on the back of this form which are subject to Missouri Department of Elementary and Secondary Education confirmation, should be taken into consideration before identifying the building(s) listed in "D" as needing a Program Improvement Plan. If documentation provided by the LEA is not approved by the SEA, a Program Improvement Plan will be submitted in accordance with the requirement.

OPTIONAL

At the request of the LEA, technical assistance in the development and implementation of the Program Improvement Plan may be provided by SEA staff.

Parents of participating children, school staff and the local educational agency jointly agree that the Missouri Department of Elementary and Secondary Education, Chapter 1 Section, be a provider of technical assistance in developing and implementing the local Program Improvement Plan.

C. LEA Definition of Substantial Progress: _____

_____

_____

_____

_____

D. Building Level Analysis.

SUBJECT AREA	GRADE SPAN

(Grade 2-12)

Method Chosen
☐ Mean
☐ Median

Local Conditions (Attach Separate Sheet of Explanation)
1. Student Mobility
2. Extent of Educational Deprivation
3. Difficulties Dealing With Secondary School Programs
4. Other Indicators of Success
5. Evaluation Design

APPENDIX D (CONT'D)

All Chapter 1 BUILDINGS	Less Than +1 NCE By Method Chosen Above	OUTCOME 1 ACHIEVED		OUTCOME 2 ACHIEVED		LOCAL CONDITIONS	✓ District Position (See 'B')			For SEA Use Only	
							Position	Position	Optional	Program Improvement	
		Yes	No	Yes	No	*(By Number)*	1	2	*Technical Assistance*	Yes	No

I certify that the above analysis relative to Chapter 1 Program Improvement is accurate, correct, and in accordance with Section 1021 of P.L. 100-297. Supporting documentation is being maintained in the district office.

_____ _____

LEA Signature of Authorized Representative — Date

_____ _____ _____ _____

Reviewed By SEA Chapter 1 Supervisor — Date — Reviewed By SEA Chapter 1 Director/Assistant Director — Date